Terms

Reliability

Reliability is about the consistency of a measure

Test-r-etest	The consistency of a measure across time: do you get the same results when you repeat the measurement?
Interrater	The consistency of a measure across raters or observers: do you get the same results when different people conduct the same measurement?
Internal consis-tency	The consistency of the measur-ement itself: do you get the same results from different parts of a test that are designed to measure the same thing?
Ensuring reliability	Apply your methods consis-tently, Standardize the conditions of your research

Validity

validity is about the accuracy of a measure

Construct	The adherence of a measure to existing theory and knowledge of the concept being measured.
Content	The extent to which the measurement covers all aspects of the concept being measured.
Criterion	The extent to which the result of a measure corresponds to other valid measures of the same concept.
Ensuring validity	Choose appropriate methods of measurement, Use appropriate sampling methods to select your subjects

Quantitative Data

is the process of collecting and analyzing numerical data. It can be used to find patterns and averages, make predictions, test causal relationships, and generalize results to wider populations.

Research Methods:

descri-ptive research	you simply seek an overall summary of your study variables.
correl-ational research	you investigate relationships between your study variables
experi-mental research	you systematically examine whether there is a cause-and-effect relationship between variables.
Advant-ages:	Replication, Direct comparison of results, Large Samples, Hypothesis testing
Disadv-antages:	Superficiality, Narrow focus, Structural bias, Lack of context

Qualitative Data

Qualitative research involves collecting and analyzing non-numerical data (e.g., text, video, or audio) to understand concepts, opinions, or experiences. It can be used to gather in-depth insights into a problem or generate new ideas for research.

Research Methods:

Observ-ations:	recording what you have seen, heard, or encountered in detailed field notes.
Interv-iews:	personally asking people questions in one-on-one conversations.
Focus groups:	asking questions and generating discussion among a group of people.
Surveys:	distributing questionnaires with open-ended questions.

Qualitative Data (cont)

Secondary research:	collecting existing data in the form of texts, images, audio or video recordings, etc.
Advant-ages:	Flexibility, Natural setting, Meaningful insights, Generation of new ideas
Disadvant-ages:	Unreliability, Subjectivity, Limited generalizability, Labor- intensive

Descriptive Statistics

summarize and organize characteristics of a data set. A data set is a collection of responses or observations from a sample or entire population.

3 main types of descriptive statistics:

1. The distribution concerns the frequency of each value (Graphs).

2. The Measures of central tendency concerns the averages of the values

- Mean	To find the mean, simply add up all response values and divide the sum by the total number of responses. The total number of responses or observations is called N.
- Median	To find the median, order each response value from the smallest to the biggest. Then, the median is the number in the middle. If there are two numbers in the middle, find their mean.
- Mode	To find the mode, order your data set from lowest to highest and find the response that occurs most frequently

3. The Measures of variability or dispersion concerns how spread out the values are

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
http://crosswordcheats.com

Descriptive Statistics (cont)

- Range	To find the range, simply subtract the lowest value from the highest value. - Standard Deviation
- Standard Deviation	The standard deviation (s) is the average amount of variability in your dataset. It tells you, on average, how far each score lies from the mean. The larger the standard deviation, the more variable the data set is.
- Variance	The variance (s2)is the average of squared deviations from the mean. Variance reflects the degree of spread in the data set. The more spread the data, the larger the variance is in relation to the mean.

Univariate descriptive statistics

Univariate descriptive statistics focus on only one variable at a time. It's important to examine data from each variable separately using multiple measures of distribution, central tendency and spread.

Bivariate descriptive statistics

If you've collected data on more than one variable, you can use bivariate or multivariate descriptive statistics to explore whether there are relationships between them. In bivariate analysis, you simultaneously study the frequency and variability of two variables to see if they vary together. You can also compare the central tendency of the two variables before performing further statistical tests.

Multivariate analysis

is the same as bivariate analysis but with more than two variables.

Contingency table

Descriptive Statistics (cont)

In a contingency table, each cell represents the intersection of two variables. Usually, an independent variable (e.g., gender) appears along the vertical axis and a dependent one appears along the horizontal axis (e.g., activities). You read "across" the table to see how the independent and dependent variables relate to each other.

Scatter plots

A scatter plot is a chart that shows you the relationship between two or three variables. It's a visual representation of the strength of a relationship.

In a scatter plot, you plot one variable along the x-axis and another one along the y-axis. Each data point is represented by a point in the chart.

Inferential Statistics

help you come to conclusions and make predictions based on your data, to understand the larger population from which the sample is taken. it's important to use random and unbiased sampling methods. If your sample isn't representative of your population, then you can't make valid statistical inferences.

Inferential statistics have two main uses:

• making estimates about populations (for example, the mean SAT score of all 11th graders in the US).

• testing hypotheses to draw conclusions about populations (for example, the relationship between SAT scores and family income).

Sampling error

Since the size of a sample is always smaller than the size of the population, some of the population isn't captured by sample data. This creates sampling error, which is the difference between the true population values (called parameters) and the measured sample values (called statistics).

Inferential Statistics (cont)

two important types of estimates you can make about the population

point estimate	is a single value estimate of a parameter. For instance, a sample mean is a point estimate of a population mean.
interval estimate	gives you a range of values where the parameter is expected to lie. A confidence interval is the most common type of interval estimate.
- confidence interval	uses the variability around a statistic to come up with an interval estimate for a parameter. Confidence intervals are useful for estimating parameters because they take sampling error into account. confidence interval tells you the uncertainty of the point estimate confidence level tells you the probability (in percentage) of the interval containing the parameter estimate if you repeat the study again A 95% confidence interval means that if you repeat your study with a new sample in exactly the same way 100 times, you can expect your estimate to lie within the specified range of values 95 times.

Hypothesis Testing

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
http://crosswordcheats.com

Inferential Statistics (cont)

is a formal process of statistical analysis using inferential statistics. The goal of hypothesis testing is to compare populations or assess relationships between variables using samples.

Parametric tests make assumptions that include the following:

• the population that the sample comes from follows a normal distribution of scores

• the sample size is large enough to represent the population

• the variances, a measure of spread, of each group being compared are similar

Non-parametric tests are called “distribution-free tests” because they don’t assume anything about the distribution of the population data.

Comparison tests

assess whether there are differences in means, medians or rankings of scores of two or more groups

T-test, Anova, Mood´s median, Wolcoxon signed- rank, Mann-Whitnes U, Krustal-Wallis H

Correlation tests

Correlation tests determine the extent to which two variables are associated.

Pearson´s r, Spearman´s r, Chi square test of independence

Regression tests

Inferential Statistics (cont)

Regression tests demonstrate whether changes in predictor variables cause changes in an outcome variable. You can decide which regression test to use based on the number and types of variables you have as predictors and outcomes. Most of the commonly used regression tests are parametric. If your data is not normally distributed, you can perform data transformations.

Simple linear regression, Multiple linear regression, Logistic regression, Nominal regression, Ordinal regression

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
http://crosswordcheats.com